BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

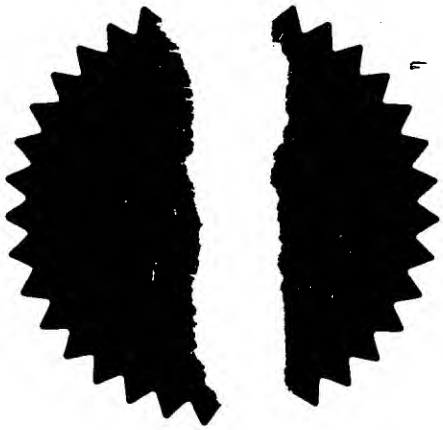
#14

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



Signed Andrew Gregg

Dated 6 FEB 2002

Patents Act 1977
(Rule 16)

Patent Office

1/77

RECD98 E414032-2 002882
P01/7700 0.00 - 9828377.3

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1.	Your reference	**SCB/51935/000**
2.	Patent application number *(The Patent Office will fill in this part)*	**9828377.3**
3.	Full name, address and postcode of the or of each applicant *(underline all surnames)*	**JANSSEN PHARMACEUTICA N.V.** **TURNHOUTSEWEG 30** **B-2340 BEERSE** **BELGIUM**
	Patents ADP number *(if you know it)*	53193001
	If the applicant is a corporate body, give the country/state of its incorporation	**BELGIUM**
4.	Title of the invention	**VASCULAR ENDOTHELIAL GROWTH FACTOR-E**
5.	Name of your agent *(if you have one)*	**BOULT WADE TENNANT** **27 FURNIVAL STREET** **LONDON** **EC4A 1PQ**
	"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*	
	Patents ADP number *(if you know it)*	**42001** ✓

6.	If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number	Country	Priority application number (if you know it)	Date of filing *(day/month/year)*

7.	If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	Number of earlier application	Date of filing *(day / month / year)*

8.	Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:* *a) any applicant named in part 3 is not an inventor, or* *b) there is an inventor who is not named as an applicant, or* *c) any named applicant is a corporate body.* *See note (d))*	**YES**

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	**0**
Description	**14**
Claim(s)	**4**
Abstract	**0**
Drawing(s)	**12** +12

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(Please specify)*	

11. I/We request the grant of a patent on the basis of this application.

Signature Date

22 December 1998

12. Name and daytime telephone number of person to contact in the United Kingdom

COLM D. MURPHY
0171 404 5921

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents ct 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 01645 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

VASCULAR ENDOTHELIAL GROWTH FACTOR-E

The present invention is concerned with a novel vascular endothelial growth factor (VEGF) herein designated "VEGF-E", and characterisation of the nucleic acid and amino acid sequences of VEGF-E.

Angiogenesis involves formation and proliferation of new blood vessels, and is an essential physiological process for normal growth and development of tissues in, for example, embryonic development, tissue regeneration and organ and tissue repair. Angiogenesis also features in the growth of human cancers which require continuous stimulation of blood vessel growth. Abnormal angiogenesis is associated with other diseases such as rheumatoid arthritis and psoriasis.

Capillary vessels consist of endothelial cells which carry the genetic information necessary to proliferate to form capillary networks. Angiogenic molecules which can initiate this process have previously been characterised. A highly selective mitogen for vascular endothelial cells is vascular endothelial growth factor (VEGF) (Ferrara *et al.*, "Vascular Endothelial Growth Factor: Basic Biology and Clinical Implications". Regulation of angiogenesis, by I.D. Goldberg and E.M. Rosen 1997 Bikhanser Vertag Basle/Switzerland). VEGF is a potent vasoactive protein which is comprised of a glycosylated cationic 46-49 kd dimer having two 24 kd subunits. It is inactivated by sulfhydryl reducing agents and is resistant to acidic pH and to heating and binds to immobilised heparin.

VEGF has four different forms of 121, 165, 189 and 206 amino acids due to alternative splicing. VEGF121 and VEGF165 are soluble and are capable of promoting angiogenesis, whereas VEGF189 and VEGF206 are bound to heparin containing proteoglycans in the cell surface. The temporal and spatial expression of VEGF has been correlated with physiological proliferation of the blood vessels (Gajdusek, C.M., and Carbon, S.J., Cell Physiol., 139:570-579, (1989)); McNeil, P.L., Muthukrishnan, L., Warder, E., D'Amore, P.A., J. Cell. Biol., 109:811-822, (1989)). Its high affinity binding sites are localized only on endothelial cells in tissue sections (Jakeman, L.B., *et al.*, Clin. Invest. 89:244-253, (1989)). The growth factor can be isolated from pituitary cells and several tumor cell lines, and has been implicated in some human gliomas (Plate, K.H. Nature 359:845-848, (1992)). The inhibition of VEGF function by anti-VEGF monoclonal antibodies was shown to inhibit tumor growth in immune-deficient mice (Kim, K.J., Nature 362:841-844, (1993)).

The present inventors have now identified a further vascular endothelial growth factor, designated herein as "VEGF-E", and the nucleic acid sequence encoding it, which has potentially significant benefits for the treatment of tumours.

Therefore, according to a first aspect of the present invention there is provided a nucleic acid molecule encoding a VEGF-E protein or a functional equvalent, derivative or bioprecursor thereof, said protein comprising the amino acid sequence illustrated in Figure 2 or 4. Preferably, the nucleic acid molecule is a DNA and even more preferably a cDNA molecule.

Also provided by this aspect of the present invention is a nucleic acid molecule such as an antisense molecule capable of hybridising to the nucleic acid molecules according to the invention under high stringency conditions.

Stringency of hybridisation as used herein refers to conditions under which polynucleic acids are stable. The stability of hybrids is reflected in the melting temperature (Tm) of the hybrids. Tm can be approximated by the formula:

$$81.5°C+16.6(\log_{10}[Na^+]+0.41\ (\%G\&C)-6001/l$$

wherein l is the length of the hybrids in nucleotides. Tm decreases approximately by 1-1.5°C with every 1% decrease in sequence homology.

The nucleic acid capable of hybridising to nucleic acid molecules according to the invention will generally be at least 70%, preferably at least 80 or 90% and more preferably at least 95% homologous to the nucleotide sequences according to the invention.

The present invention also comprises within its scope proteins or polypeptides encoded by the nucleic acid molecules according to the invention or a functional equivalent, derivative or bioprecursor thereof.

Therefore, according to a further aspect of the present invention, there is provided a VEGF-E protein, or a functional equivalent, derivative or bioprecursor thereof, having an amino acid sequence as illustrated in Figure 2 or 4. A further aspect of the invention comprises a VEGF-E protein, or a functional

equivalent, derivative or bioprecursor thereof, encoded by a nucleic acid molecule according to the invention. Preferably, the VEGF-E protein encoded by said nucleic acid molecule comprises an amino acid sequence as illustrated in Figure 2 or 4.

The DNA molecules according to the invention may, advantageously, be included in a suitable expression vector to express VEGF-E encoded therefrom in a suitable host.

An expression vector according to the invention includes a vector having a nucleic acid according to the invention operably linked to regulatory sequences, such as promoter regions, that are capable of effecting expression of said DNA fragments. The term "operably linked" refers to a juxta position wherein the components described are in a relationship permitting them to function in their intended manner. Such vectors may be transformed into a suitable host cell to provide for expression of a polypeptide according to the invention. Thus, in a further aspect, the invention provides a process for preparing polypeptides according to the invention which comprises cultivating a host cell, transformed or transfected with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the polypeptides, and recovering the expressed polypeptides.

The vectors may be, for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of said nucleotide and optionally a regulator of the promoter.

The vectors may contain one or more selectable markers, such as, for example, ampicillin resistance.

Regulatory elements required for expression include promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. For example, a bacterial expression vector may include a promoter such as the lac promoter and for transcription initiation the Shine-Dalgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector may include a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or assembled from the sequences described by methods well known in the art.

Nucleic acid molecules according to the invention may be inserted into the vectors described in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense nucleic acids may be produced by synthetic means.

In accordance with the present invention, a defined nucleic acid includes not only the identical nucleic acid but also any minor base variations including in particular, substitutions in bases which result in a synonymous codon (a different codon specifying the same amino acid residue) due to the degenerate code in conservative amino acid substitutions. The term "nucleic acid sequence" also includes the complementary sequence to any single stranded sequence given regarding base variations.

The present invention also advantageously provides nucleic acid sequences of at least approximately 10 contiguous nucleotides of a nucleic acid according to the invention and preferably from 10 to 50 nucleotides. These sequences may, advantageously be used as probes or primers to initiate replication, or the like. Such nucleic acid sequences may be produced according to techniques well known in the art, such as by recombinant or synthetic means. They may also be used in diagnostic kits or the like for detecting the presence of a nucleic acid according to the invention. These tests generally comprise contacting the probe with the sample under hybridising conditions and detecting for the presence of any duplex or triplex formation between the probe and any nucleic acid in the sample.

The nucleic acid sequences according to this aspect of the present invention comprises the sequences of nucleotides designated herein as VEGFE 1-10, illustrated in Figure 5.

According to the present invention these probes may be anchored to a solid support. Preferably, they are present on an array so that multiple probes can simultaneously hybridize to a single biological sample. The probes can be spotted onto the array or synthesised *in situ* on the array. (See Lockhart *et al.*, Nature Biotechnology, vol. 14, December 1996 "Expression monitoring by hybridisation to high density oligonucleotide arrays". A single array can contain more than 100, 500 or even 1,000 different probes in discrete locations.

The nucleic acid sequences, according to the invention

may be produced using such recombinant or synthetic means, such as for example using PCR cloning mechanisms which generally involve making a pair of primers, which may be from approximately 10 to 50 nucleotides to a region of the gene which is desired to be cloned, bringing the primers into contact with mRNA, cDNA, or genomic DNA from a human cell, performing a polymerase chain reaction under conditions which bring about amplification of the desired region, isolating the amplified region or fragment and recovering the amplified DNA. Generally, such techniques as defined herein are well known in the art, such as described in Sambrook *et al* (Molecular Cloning: a Laboratory Manual, 1989).

The nucleic acids or oligonucleotides according to the invention may carry a revealing label. Suitable labels include radioisotopes such as ^{32}P or ^{35}S, enzyme labels or other protein labels such as biotin or fluorescent markers. Such labels may be added to the nucleic acids or oligonucleotides of the invention and may be detected using known techniques *per se*.

The protein according to the invention includes all possible amino acid variants encoded by the nucleic acid molecule according to the invention including a polypeptide encoded by said molecule and having conservative amino acid changes. Proteins or polypeptides according to the invention further include variants of such sequences, including naturally occurring allelic variants which are substantially homologous to said proteins or polypeptides. In this context, substantial homology is regarded as a sequence which has at least 70%, preferably 80 or 90% amino acid homology with the

proteins or polypeptides encoded by the nucleic acid molecules according to the invention.

The nucleic acid or protein according to the invention may be used as a medicament or in the preparation of a medicament for treating cancer or other diseases or conditions associated with expression of VEGF-E protein.

Advantageously, the nucleic acid molecule or the protein according to the invention may be provided in a pharmaceutical composition together with a pharmacologically acceptable carrier, diluent or excipient therefor.

The present invention is further directed to inhibiting VEGF2 *in vivo* by the use of antisense technology. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the mature protein sequence, which encodes for the protein of the present invention, is used to design an antisense RNA oligonucleotide of from 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple-helix - see Lee *et al.* Nucl. Acids Res., 6:3073 (1979); Cooney *et al.*, Science, 241:456 (1988); and Dervan *et al.*, Science, 251: 1360 (1991), thereby preventing transcription and the production of VEGF2. The antisense RNA oligonucleotide hybridises to the mRNA *in vivo* and blocks translation of an mRNA molecule into the VEGF2 (antisense - Okano, J. Neurochem., 56:560 (1991); Oligodeoxynucleotides as

Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)).

Alternatively, the oligonucleotide described above can be delivered to cells by procedures in the art such that the anti-sense RNA or DNA may be expressed *in vivo* to inhibit production of VEGF-E in the manner described above.

Antisense constructs to VEGF-E, therefore, may inhibit the angiogenic activity of the VEGF-E and prevent the further growth or even regress solid tumours, since angiogenesis and neovascularization are essential steps in solid tumour growth. These antisense constructs may also be used to treat rheumatoid arthritis, psoriasis and diabetic retinopathy which are all characterized by abnormal angiogenesis.

A further aspect of the invention provides a host cell or organism, transformed or transfected with an expression vector according to the invention. The host cell or organism may advantageously be used in a method of producing VEGF-E, which comprises recovering any expressed VEGF-E from the host or organism transformed or transfected with the expression vector.

According to a further aspect of the invention there is also provided a transgenic cell, tissue or organism comprising a transgene capable of expressing VEGF-E protein according to the invention. The term "transgene capable of expression" as used herein means a suitable nucleic acid sequence which leads to expression of VEGF-E or proteins having the same function and/or activity. The transgene, may include, for example, genomic nucleic acid isolated from human

cells or synthetic nucleic acid, including DNA integrated into the genome or in an extrachromosomal state. Preferably, the transgene comprises the nucleic acid sequence encoding the proteins according to the invention as described herein, or a functional fragment of said nucleic acid. A functional fragment of said nucleic acid should be taken to mean a fragment of the gene comprising said nucleic acid coding for the proteins according to the invention or a functional equivalent, derivative or a non-functional derivative such as a dominant negative mutant, or bioprecursor of said proteins. For example, it would be readily apparent to persons skilled in the art that nucleotide substitutions or deletions may be used using routine techniques, which do not affect the protein sequence encoded by said nucleic acid, or which encode a functional protein according to the invention.

VEGF-E protein expressed by said transgenic cell, tissue or organism or a functional equivalent or bioprecursor of said protein also form part of the present invention.

Antibodies to the protein or polypeptide of the present invention may, advantageously, be prepared by techniques which are known in the art. For example, polyclonal antibodies may be prepared by inoculating a host animal, such as a mouse, with the polypeptide according to the invention or an epitope thereof and recovering immune serum. Monoclonal antibodies may be prepared according to known techniques such as described by Kohler R. and Milstein C., Nature (1975) 256, 495-497.

Antibodies according to the invention may also be used in a method of detecting for the presence of a polypeptide according to the invention, which method comprises reacting the antibody with a sample and identifying any protein bound to said antibody. A kit may also be provided for performing said method which comprises an antibody according to the invention and means for reacting the antibody with said sample.

Proteins which interact with the polypeptide of the invention may be identified by investigating protein-protein interactions using the two-hybrid vector system first proposed by Chien *et al* (1991).

This technique is based on functional reconstitution in vivo of a transcription factor which activates a reporter gene. More particularly the technique comprises providing an appropriate host cell with a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA binding domain and an activating domain, expressing in the host cell a first hybrid DNA sequence encoding a first fusion of a fragment or all of a nucleic acid sequence according to the invention and either said DNA binding domain or said activating domain of the transcription factor, expressing in the host at least one second hybrid DNA sequence, such as a library or the like, encoding putative binding proteins to be investigated together with the DNA binding or activating domain of the transcription factor which is not incorporated in the first fusion; detecting any binding of the proteins to be investigated with a protein according to the invention by detecting for the presence of any reporter gene product in the host cell; optionally isolating second

hybrid DNA sequences encoding the binding protein.

An example of such a technique utilises the GAL4 protein in yeast. GAL4 is a transcriptional activator of galactose metabolism in yeast and has a separate domain for binding to activators upstream of the galactose metabolising genes as well as a protein binding domain. Nucleotide vectors may be constructed, one of which comprises the nucleotide residues encoding the DNA binding domain of GAL4. These binding domain residues may be fused to a known protein encoding sequence, such as for example the nucleic acids according to the invention. The other vector comprises the residues encoding the protein binding domain of GAL4. These residues are fused to residues encoding a test protein. Any interaction between polypeptides encoded by the nucleic acid according to the invention and the protein to be tested leads to transcriptional activation of a reporter molecule in a GAL-4 transcription deficient yeast cell into which the vectors have been transformed. Preferably, a reporter molecule such as ß-galactosidase is activated upon restoration of transcription of the yeast galactose metabolism genes.

Advantageously, the antibody according to the invention may also be used as a medicament or in the preparation of a medicament for treating tumours or other diseases associated with expression of VEGF-E. The invention also further provides a pharmaceutical composition comprising said antibody together with a pharmaceutically acceptable carrier diluent or excipient therefor.

A further aspect of the present invention also

provides a method of identifying VEGF-E in a sample, which method comprises contacting said sample with an antibody according to the invention and monitoring for any hybridisation of any proteins to said antibody. A kit for identifying the presence of VEGF-E in a sample is also provided comprising an antibody according to the invention and means for contacting said antibody with said sample.

The invention may be more clearly understood with reference to the accompanying example, which is purely exemplary, with reference to the accompanying drawings, wherein:

Figure 1: is a nucleotide sequence coding for a partial VEGF-E protein according to the invention.

Figure 2: is an illustration of amino acid sequence of the nucleic acid sequence of Figure 1.

Figure 3: is an illustration of a nucleotide sequence encoding VEGF-E protein according to the invention.

Figure 4: is an illustration of the amino acid sequence of the nucleic acid sequence of Figure 3.

Figure 5: depicts the nucleic acid sequences of the first 18 human EST clones obtained from the BLAST search of the LifSeq™ database.

Figure 6: depicts the nucleotide sequences of 50 human EST clones obtained from the proprietary

LifeSeq™ database.

Figure 7: is an illustration of the nucleotide sequences utilised as primers to identify the sequence of the gene coding for VEGF-E.

EXAMPLE 1

A BLAST (Basic Local Alignment Search Tool; Altschul *et al.*, 1990 J. Mol. Biol. 215, 403-410) search was performed in the propriety LifeSeq™ human EST database (Incyte Pharmaceuticals, Inc., Palo Alto, CA, USA). BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologues. While it is useful for matches which do not contain gaps, it is inappropriate for performing motif-style searching. The fundamental unit of BLAST algorithm output is the High-scoring Segment Pair (HSP).

Eighteen human EST clones (Figure 5) with high similarity to the previously identified VEGF proteins were identified and a further fifty EST clones (Figure 6) were identified using these sequences as query sequences, allowing us to deduce the putative sequence for the new VEGF-E protein. The sequences obtained were compared to known sequences to determine regions of homology and to identify the sequence as a novel VEGF-E protein. Using the DNA sequence information in the databases we were able to prepare suitable primers having the sequences of VEGFE 1-10 illustrated in Figure 7 for use in subsequent RACE experiments to obtain the complete DNA sequence for the VEGF-E gene.

CLAIMS

1. A nucleic acid molecule encoding a VEGF-E protein or a functional equivalent derivative or bioprecursor thereof, said protein comprising the amino acid sequence illustrated in Figures 2 or 4.

2. A nucleic acid molecule according to claim 1 wherein said nucleic acid is a DNA molecule.

3. A nucleic acid molecule according to claim 1 or 2 wherein said nucleic acid is a cDNA molecule.

4. A nucleic acid molecule according to any of claims 1 to 3 comprising the nucleotide sequence illustrated in Figure 1 or 3.

5. A nucleic acid molecule capable of hybridising to a molecule according to any of claims 1 to 4 under high stringency conditions.

6. A VEGF-E protein, or a functional equivalent, derivative or bioprecursor thereof, having the amino acid sequence illustrated in Figure 2 or 4.

7. A VEGF-E protein, or a functional equivalent, derivative or bioprecursor thereof, encoded by a nucleic acid molecule according to any of claims 1 to 4.

8. A protein according to claim 7, which comprises the amino acid sequence illustrated in Figure 2 or 4.

9. An expression vector comprising a nucleic acid molecule according to any of claims 1 to 4.

10. An expression vector according to claim 9 further comprising a nucleotide sequence encoding a reporter molecule.

11. A nucleic acid molecule according to any of claims 1 to 5 for use as a medicament.

12. Use of a nucleic acid molecule according to any of claims 1 to 5 in the preparation of a medicament for inhibiting angiogenic activity and formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair or for treating cancer or rheumatoid arthritis or psoriasis or diabetic retinopathy.

13. A pharmaceutical composition comprising a nucleic acid molecule or a protein according to any of claims 1 to 5 or 6 to 8 respectively, together with a pharmaceutically acceptable carrier, diluent or excipient therefor.

14. A host cell or organism transformed or transfected with an expression vector according to claim 9 or 10.

15. A transgenic cell, tissue or organism comprising a transgene capable of expressing a VEGF-E protein according to any of claims 6 to 8.

16. A process for producing a VEGF-E protein according to any of claims 6 to 8, said process comprising transforming a host cell or organism with an expression vector according to claim 9 and 10, and recovering the expressed protein from said host cell or organism.

17. An antibody capable of binding to a protein according to any of claims 6 to 8, which is preferably a monoclonal antibody.

18. An antibody according to claim 17 for use as a medicament.

19. Use of an antibody according to claim 17 in the preparation of a medicament for inhibiting angiogenic activity and formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair or for treating cancer or rheumatoid arthritis or psoriasis or diabetic retinopathy.

20. A pharmaceutical composition comprising an antibody according to claim 17 together with a pharmaceutically acceptable carrier diluent or excipient therefor.

21. A method of identifying VEGF-E in a sample which method comprises contacting said sample with an antibody according to claim 17 and monitoring for binding of any protein to said antibody.

22. A kit for identifying the presence of VEGF-E in a sample which comprises an antibody according to claim 17 and means for contacting said antibody with said sample.

23. A method of identifying compounds which inhibit angiogenesis which method comprises providing a host cell or organism according to claim 14 or a transgenic

cell, tissue or organism according to claim 15, contacting a test compound with said cell, tissue or organism and monitoring for the presence or absence either of said reporter molecule or VEGF-E.

24. A compound identifiable according to the method of claim 23.

25. A compound according to claim 24 for use as a medicament.

26. Use of a compound according to claim 24 in the preparation of a medicament for inhibiting angiogenic activity and formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair or for treating cancer, rheumatoid arthritis, psoriasis or diabetic retinopathy.

27. A nucleic acid sequence comprising the nucleotide sequence of any of the sequences identified in Figure 6 or 7.

28. An expression vector comprising a nucleic acid sequence according to claim 27.

29. A host cell transformed or transfected with an expression vector according to claim 28.

30. A method for producing a polypeptide, said method comprising the steps of:

a) culturing the host cell of claim 29 under conditions suitable for expression of the peptide; and

b) recovering the polypeptide from the host cell culture.

```
+3                                          M  N  I  F  L  L  N  L  L  T  E  E  V  R  L  Y
                                            ]-----------------------------------------------
1   AGGAAATCAA ATTAGGATAA GATTTGTATC TGATGAATAT TTTCCTTCTG AACCTTCTAA CAGAGGAGGT AAGATTATAC
    TCCTTTAGTT TAATCCTATT CTAAACATAG ACTACTTATA AAAGGAAGAC TTGGAAGATT GTCTCCTCCA TTCTAATATG
.....
+3  S  C  T  P  R  N  P  S  V  S  I  R  E  E  L  K  R  T  D  T  I  F  W  P  G  C  L
    ------------------------]-----------------------------------------------------------
81  AGCTGCACAC CTCGTAACTT CTCAGTGTCC ATAAGGGAAG AACTAAAGAG AACCGATACC ATTTTCTGGC CAGGTTGTCT
    TCGACGTGTG GAGCATTGAA GAGTCACAGG TATTCCCTTC TTGATTTCTC TTGGCTATGG TAAAAGACCG GTCCAACAGA
-2                                                  <-------------------------------------
.....
+3    L  V  K  R  C  G  G  N  C  A  C  C  L  H  N  C  N  E  C  Q  C  V  P  S  K  V
    -----------------------------------------------------------------------------------
161 CCTGGTTAAA CGCTGTGGTG GGAACTGTGC CTGTTGTCTC CACAATTGCA ATGAATGTCA ATGTGTCCCA AGCAAAGTTA
    GGACCAATTT GCGACACCAC CCTTGACACG GACAACAGAG GTGTTAACGT TACTTACAGT TACACAGGGT TCGTTTCAAT
-2  -----------------------------------------------------------------------------------
.....
+3  T  K  K  Y  H  E  V  L  Q  L  R  P  K  T  G  V  R  G  L  H  K  S  L  T  D  V  A
    -----------------------------------------------------------------------------------
+1                                                V  S  G  D  C  T  N  H  S  P  T  W  P
                                                  ]-------------------------------------
241 CTAAAAAATA CCACGAGGTC CTTCAGTTGA GACCAAAGAC CGGTGTCAGG GGATTGCACA AATCACTCAC CGACGTGGCC
    GATTTTTTAT GGTGCTCCAG GAAGTCAACT CTGGTTTCTG GCCACAGTCC CCTAACGTGT TTAGTGAGTG GCTGCACCGG
-2  ---------------------------------------------------------[
.....
+3  L  E  H  H  E  E  C  D  C  V  C  R  G  S  T  G  G
    ----------------------------------------------->
+2                                   V  Q  R  E  H  R  R  I  A  A  S  P  P  A  A  L  A
                                     ]-----------------------------------------------------
+1  W  S  T  M  R  S  V  T  V  C  A  E  G  A  Q  E  D  S  R  I  T  T  S  S  S  C
    -----------------------------------------------------------------------------------
321 CTGGAGCACC ATGAGGAGTG TGACTGTGTG TGCAGAGGGA GCACAGGAGG ATAGCCGCAT CACCACCAGC AGCTCTTGCC
    GACCTCGTGG TACTCCTCAC ACTGACACAC ACGTCTCCCT CGTGTCCTCC TATCGGCGTA GTGGTGGTCG TCGAGAACGG
.....
+2  Q  S  C  A  V  Q  W  L  I  L  L  E  N  V  C  V  I  S  I  L  N  L  S  C  L  L  Q
    -----------------------------------------------------------------------------------
+1 P  E  L  C  S  A  V  A  D  S  I  R  E  R  M  R  Y  L  H  P
    ------------------------------------------------------------>
401 CAGAGCTGTG CAGTGCAGTG GCTGATTCTA TTAGAGAACG TATGCGTTAT CTCCATCCTT AATCTCAGTT GTTTGCTTCA
    GTCTCGACAC GTCACGTCAC CGACTAAGAT AATCTCTTGC ATACGCAATA GAGGTAGGAA TTAGAGTCAA CAAACGAAGT
.....
+2  G  P  F  I  F  R  I  Y  S  A  F
    ------------------------------->
481 AGGACCTTTC ATCTTCAGGA TTTACAGTGC ATTCTGAAAG AGGAGACATC AAACAGAATT AGGAGTTGTG CAACAGCTCT
    TCCTGGAAAG TAGAAGTCCT AAATGTCACG TAAGACTTTC TCCTCTGTAG TTTGTCTTAA TCCTCAACAC GTTGTCGAGA
.....
561 TTTGAGAGGA CGCCTAAAGG ACAGGAGAAA AGGTCTTCAA TCGTGGAAAG AAAATTAAAT GTTGTATTAA ATAGATCACC
    AAACTCTCCT CCGGATTTCC TGTCCTCTTT TCCAGAAGTT AGCACCTTTC TTTTAATTTA CAACATAATT TATCTAGTGG
.....
641 AGCTAGTTTC AGAGTTACCA TGTACGTATT CCACTAGCTG GGTTCTGTAT TTCAGTTCTT TCGATACGGC TTAGGGTAAT
    TCGATCAAAG TCTCAATGGT ACATGCATAA GGTGATCGAC CCAAGACATA AAGTCAAGAA AGCTATGCCG AATCCCATTA
.....
721 GTCAGTACAG GAAAAAAACT GTGCAAGTGA GCACCTGATT CCGTTGCCTT GGCTTAACTC TAAAGCTCCA TGTCCTGGGC
    CAGTCATGTC CTTTTTTTGA CACGTTCACT CGTGGACTAA GGCAACGGAA CCGAATTGAG ATTTCGAGGT ACAGGACCCG
.....
801 CTAAAATCGT ATAAAATCTG GA
    GATTTTAGCA TATTTTAGAC CT
    .....
```

Fig 1

```
  1  MNIFLLNLLT EEVRLYSCTP RNFSVSIREE LKRTDTIFWP GCLLVKRCGG
 51  NCACCLHNCN ECQCVPSKVT KKYHEVLQLR PKTGVRGLHK SLTDVALEHH
101  EECDCVCRGS TGG
```

Fig 2

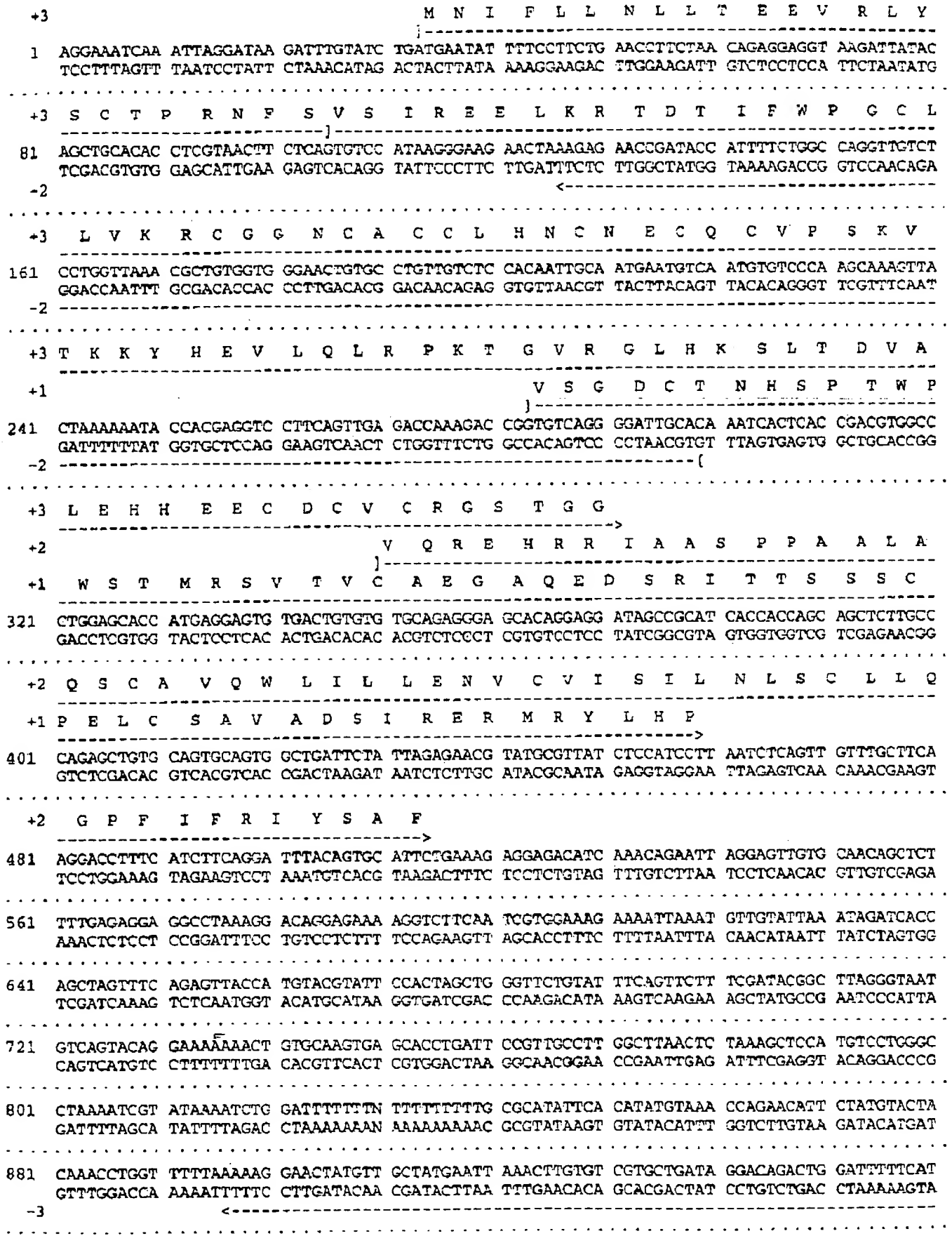

+3 M N I F L L N L L T E E V R L Y
1 AGGAAATCAA ATTAGGATAA GATTTGTATC TGATGAATAT TTTCCTTCTG AACCTTCTAA CAGAGGAGGT AAGATTATAC
TCCTTTAGTT TAATCCTATT CTAAACATAG ACTACTTATA AAAGGAAGAC TTGGAAGATT GTCTCCTCCA TTCTAATATG
+3 S C T P R N F S V S I R E E L K R T D T I F W P G C L
81 AGCTGCACAC CTCGTAACTT CTCAGTGTCC ATAAGGGAAG AACTAAAGAG AACCGATACC ATTTTCTGGC CAGGTTGTCT
TCGACGTGTG GAGCATTGAA GAGTCACAGG TATTCCCTTC TTGATTTCTC TTGGCTATGG TAAAAGACCG GTCCAACAGA
-2
+3 L V K R C G G N C A C C L H N C N E C Q C V P S K V
161 CCTGGTTAAA CGCTGTGGTG GGAACTGTGC CTGTTGTCTC CACAATTGCA ATGAATGTCA ATGTGTCCCA AGCAAAGTTA
GGACCAATTT GCGACACCAC CCTTGACACG GACAACAGAG GTGTTAACGT TACTTACAGT TACACAGGGT TCGTTTCAAT
-2
+3 T K K Y H E V L Q L R P K T G V R G L H K S L T D V A
+1 V S G D C T N H S P T W P
241 CTAAAAAATA CCACGAGGTC CTTCAGTTGA GACCAAAGAC CGGTGTCAGG GGATTGCACA AATCACTCAC CGACGTGGCC
GATTTTTTAT GGTGCTCCAG GAAGTCAACT CTGGTTTCTG GCCACAGTCC CCTAACGTGT TTAGTGAGTG GCTGCACCGG
-2
+3 L E H H E E C D C V C R G S T G G
+2 V Q R E H R R I A A S P P A A L A
+1 W S T M R S V T V C A E G A Q E D S R I T T S S S C
321 CTGGAGCACC ATGAGGAGTG TGACTGTGTG TGCAGAGGGA GCACAGGAGG ATAGCCGCAT CACCACCAGC AGCTCTTGCC
GACCTCGTGG TACTCCTCAC ACTGACACAC ACGTCTCCCT CGTGTCCTCC TATCGGCGTA GTGGTGGTCG TCGAGAACGG
+2 Q S C A V Q W L I L L E N V C V I S I L N L S C L L Q
+1 P E L C S A V A D S I R E R M R Y L H P
401 CAGAGCTGTG CAGTGCAGTG GCTGATTCTA TTAGAGAACG TATGCGTTAT CTCCATCCTT AATCTCAGTT GTTTGCTTCA
GTCTCGACAC GTCACGTCAC CGACTAAGAT AATCTCTTGC ATACGCAATA GAGGTAGGAA TTAGAGTCAA CAAACGAAGT
+2 G P F I F R I Y S A F
481 AGGACCTTTC ATCTTCAGGA TTTACAGTGC ATTCTGAAAG AGGAGACATC AAACAGAATT AGGAGTTGTG CAACAGCTCT
TCCTGGAAAG TAGAAGTCCT AAATGTCACG TAAGACTTTC TCCTCTGTAG TTTGTCTTAA TCCTCAACAC GTTGTCGAGA
561 TTTGAGAGGA GGCCTAAAGG ACAGGAGAAA AGGTCTTCAA TCGTGGAAAG AAAATTAAAT GTTGTATTAA ATAGATCACC
AAACTCTCCT CCGGATTTCC TGTCCTCTTT TCCAGAAGTT AGCACCTTTC TTTTAATTTA CAACATAATT TATCTAGTGG
641 AGCTAGTTTC AGAGTTACCA TGTACGTATT CCACTAGCTG GGTTCTGTAT TTCAGTTCTT TCGATACGGC TTAGGGTAAT
TCGATCAAAG TCTCAATGGT ACATGCATAA GGTGATCGAC CCAAGACATA AAGTCAAGAA AGCTATGCCG AATCCCATTA
721 GTCAGTACAG GAAAAAAACT GTGCAAGTGA GCACCTGATT CCGTTGCCTT GGCTTAACTC TAAAGCTCCA TGTCCTCGGC
CAGTCATGTC CTTTTTTTGA CACGTTCACT CGTGGACTAA GGCAACGGAA CCGAATTGAG ATTTCGAGGT ACAGGACCCG
801 CTAAAATCGT ATAAAATCTG GATTTTTTTN TTTTTTTTTG CGCATATTCA CATATGTAAA CCAGAACATT CTATGTACTA
GATTTTAGCA TATTTTAGAC CTAAAAAAAN AAAAAAAAAC GCGTATAAGT GTATACATTT GGTCTTGTAA GATACATGAT
881 CAAACCTGGT TTTTAAAAAG GAACTATGTT GCTATGAATT AAACTTGTGT CGTGCTGATA GGACAGACTG GATTTTTCAT
GTTTGGACCA AAAATTTTTC CTTGATACAA CGATACTTAA TTTGAACACA GCACGACTAT CCTGTCTGAC CTAAAAAGTA
-3

Fig 3

```
 961  ATTTCTTATT AAAATTTCTG CCATTTAGAA GAAGAGAACT ACATTCATGG TTTGGAAGAG ATAAACCTGA AAAGAAGAGT
      TAAAGAATAA TTTTAAAGAC GGTAAATCTT CTTCTCTTGA TGTAAGTACC AAACCTTCTC TATTTGGACT TTTCTTCTCA
   -3 ----------------------------------------------------------------------------------

1041  GGCCTTATCT TCACTTTATC GATAAGTCAG TTTATTTGTT TCATTGTGTA CATTTTTATA TTCTCCTTTT GACATTATAA
      CCGGAATAGA AGTGAAATAG CTATTCAGTC AAATAAACAA AGTAACACAT GTAAAAATAT AAGAGGAAAA CTGTAATATT
   -3 -------------[

1121  CTGTTGGCTT TTCTAATCTT GTTAAATATA TCTATTTTTA CCAAAGGTAT TTAATATTCT TTTTTATGAC AACTTAGATC
      GACAACCGAA AAGATTAGAA CAATTTATAT AGATAAAAAT GGTTTCCATA AATTATAAGA AAAAATACTG TTGAATCTAG

1201  AACTATTTTT AGCTTGGTAA ATTTTTCTAA ACACAATTGT TATAGCCAGA GGAACAAAGA TGATATAAAA TATTGTTGCT
      TTGATAAAAA TCGAACCATT TAAAAAGATT TGTGTTAACA ATATCGGTCT CCTTGTTTCT ACTATATTTT ATAACAACGA

1281  CTGACAAAAA TACATGTATT TCATTCTCGT ATGGTGCTAG AGTTAGATTA ATCTGCATTT TAAAAAACTG AATTGGAATA
      GACTGTTTTT ATGTACATAA AGTAAGAGCA TACCACGATC TCAATCTAAT TAGACGTAAA ATTTTTTGAC TTAACCTTAT

1361  GAATTGGTAA GTTGCAAAGA CTTTTTGAAA ATAATTAAAT TATCATATCT TCCATTCCTG TTATTGGAGA TGAAAATAAA
      CTTAACCATT CAACGTTTCT GAAAAACTTT TATTAATTTA ATAGTATAGA AGGTAAGGAC AATAACCTCT ACTTTTATTT

1441  AAGCAACTTA TGAAAGTAGA CATTCAGATC CAGCCATTAC TAACCTATTC CTTTTTTGGG GAAATCTGAG CCTAGCTCAG
      TTCGTTGAAT ACTTTCATCT GTAAGTCTAG GTCGGTAATG ATTGGATAAG GAAAAAACCC CTTTAGACTC GGATCGAGTC

1521  AAAAACATAA AGCACCTTGA AAAAGACTTG GCAGCTTCCT GATAAAGCGT GCTGTGCTGT GCAGTAGGAA CACATCCTAT
      TTTTTGTATT TCGTGGAACT TTTTCTGAAC CGTCGAAGGA CTATTTCGCA CGACACGACA CGTCATCCTT GTGTAGGATA

1601  TTATTGTGAT GTTGTGGTTT TATTATCTTA AACTCTGTTC CATACACTTG TATAAATACA TGGATATTTT TATGTACAGA
      AATAACACTA CAACACCAAA ATAATAGAAT TTGAGACAAG GTATGTGAAC ATATTTATGT ACCTATAAAA ATACATGTCT

1681  AGTATGTCTC TTAACCAGTT CACTTATTGT ACCTGG
      TCATACAGAG AATTGGTCAA GTGAATAACA TGGACC
```

Fig 3 (cont'd)

```
  1  MNIFLLNLLT EEVRLYSCTP RNFSVSIREE LKRTDTIFWP GCLLVKRCGG
 51  NCACCLHNCN ECQCVPSKVT KKYHEVLQLR PKTGVRGLHK SLTDVALEHH
101  EECDCVCRGS TGG
```

Figure 4

```
>3993180H1      LUNGNON03       INCYTE
CACAGTTCACTCACCGACGTGGCCCTGGAGCACCATGAGGNGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCC
GCATCCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCAT
CCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAG
AATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCTAAAGGACAGGAGAANAGGTCTT
>3510192H1      CONCNOT01       INCYTE
TGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTT
TCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAG
GAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATTAAATAGATCACCAGCTAGTT
TCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTT
>2559870H1      ADRETUT01       INCYTE
CACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCA
TGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGGGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGC
AGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCA
TCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGA
>3979767H1      LUNGTUT08       INCYTE
GGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGC
GTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAG
ACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTG
GAAAGAANATTAAATGTTGTATTAAATAGACACCAGCT
>3980011H1      LUNGTUT08       INCYTE
GGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGC
GTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACATGCATTCTGAAAGAGGAGA
CATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGG
AAAGAAAATTAAATGTTGTATTAAATAGATCACCA
>4825396H1      BLADDIT01       INCYTE
GAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCACAATT
GCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTC
AGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGG
AGGATAGCCGCATCACCACCA
>3073703H1      BONEUNT01       INCYTE
AGAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCTCAGT
GTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAACT
GTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAG
TTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCA
>1302516H1      PLACNOT02       INCYTE
AGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACCTTCTAACAGAGGAGGTAAGATTATAC
AGCTGCACACCTCGTAACTTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCT
CCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTT
ACTAAAAAATACCACGAGGTCC
>3684109H1      HEAANOT01       INCYTE
ATTTCATCTTCAGGATTTACAGTGCATTCTGAAANAGGAGAAATCAAACANAATTAGGAGTTGTGCAACAGCTCTTTTGA
GAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAANAAAATTAAATGTTGTATTAAATAGATCACCAGCTA
GTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGGTAATGTCAG
TACAGGAAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTT
>4713188H1      BRAIHCT01       INCYTE
CAAAGTTACTAAAAAATACCACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCACTCACCG
ACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAG
CTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGT
TTGCT
>458823H1      KERANOT01       INCYTE
ANGAGTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTT
GTTTGNTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTG
CAACAGCTCTTTTGAGAGGAGGCCTAAAGGNCAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATTAA
ATAGATC
>1303909H1      PLACNOT02       INCYTE
AGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACCTTCTAACAGAGGAGGTAAGATTATAC
AGCTGCACACCTCGTAACTTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCT
CCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAG
>2739211H1      OVARNOT09       INCYTE
GTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGA
GAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATTAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACG
TATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAA
GTGAGCACCTGAT
>3325591H1      PTHYNOT03       INCYTE
TGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATT
AAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACG
GCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACCCTAAAGCNCC
ATGTCNNGGGCNAAAANCGAAAAAT
>3733565H1      SMCCNOS01       INCYTE
CCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGNAAGANGAGACATCAAACAG
AATTAGGNGTTGTGCAAAAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTNCAATCGTGGAAAGNAAATT
AAATGTTGTATNAAATNGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGNCNGTATTCAGTCT
TTCGGAACGGCTTAGGGTAATGTCAGTACAGGANAAAAACTGTGCAGTGAG
>3554223H1      SYNONOT01       INCYTE
```

file: short_est_mples

Fig 5

```
76 ATTAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGAT
   ACGGCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGGCTTAACTCTAAAG
   CTCC  TCCTGGGCCTAAAATCGTATAAAATCTGGATTTTTTTNTTTTTTTTTGCGCATATTCACATATGTAAACCAGN
79 ACATTCTATGTACNACAAACCTGGTTTTTAAAAAGGAAC
80 >4507477H1       OVARTDT01        INCYTE
81 GGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGCTAAT
82 GTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCCATGTCCTGGGCC
83 TAAAATCGTATAAAATCTGGA
84 >4163378H1       BRSTNOT32        INCYTE
85 AATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGNTCTGTATTTCAGTTCCTTTCGATACG
86 GCTTAGGGTAATGTCAGTACAGGAAAAAAGCTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCC
87 ATGTCCTGGGCCTAAAATCGTATA
```

Fig 5 (cont'd)

```
1  >2054675H1       BEPINOT01        INCYTE
   AAAGGAACTATGTTGCTATGAATTAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAATT
   TCTGCTTTAGAAGAAGAGAACTACATTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTT
4  TATCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAA
5  TCTTGTTAAATATATCTATTTTTACCAAAGGTATTTAATATTCTTTTTTA
6  >3993180H1       LUNGNON03        INCYTE
7  CACAAATCACTCACCGACGTGGCCCTGGAGCACCATGAGGNGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCC
8  GCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCAT
9  CCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAG
10 AATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCTAAAGGACAGGAGAANAGGTCTT
11 >3510192H1       CONCNOT01        INCYTE
12 TGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTT
13 TCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAG
14 GAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATTAAATAGATCACCAGCTAGTT
15 TCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTT
16 >4164633H1       BRSTNOT32        INCYTE
17 CTTGTTAAATATATCTATTTTTACCAAAGGTATTTAATATTCTTTANTTATGACAACTTAGATCAACTATTTTTAGCTTG
18 GTAAATTTTTCTAAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATG
19 TATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCA
20 AAGACTTTTTGANAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGGGGAGAAAAT
21 >2559870H1       ADRETUT01        INCYTE
22 CACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCA
23 TGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGGGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGC
24 AGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCA
25 TCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGA
26 >3817470H1       BONSTUT01        INCYTE
27 TTAAAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCATGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAA
28 AATTTCTGCCATTTAGAAGAAGAGAACTACATTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTC
29 ACTTTATCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTC
30 TAATCTGTTAAATATATCTATTTTTACCAAAGGTATTTAATATTCTTT
31 >3979767H1       LUNGTUT08        INCYTE
32 GGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGC
33 GTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAG
34 ACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTG
35 GAAAGAANATTAAATGTTGTATTAAATAGACACCAGCT
36 >3980011H1       LUNGTUT08        INCYTE
37 GGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGC
38 GTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACATGCATTCTGAAAGAGGAGA
39 CATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGG
40 AAAGAAAATTAAATGTTGTATTAAATAGATCACCA
41 >4825396H1       BLADDIT01        INCYTE
42 GAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCACAATT
43 GCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTC
44 AGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGG
45 AGGATAGCCGCATCACCACCA
46 >3073703H1       BONEUNT01        INCYTE
47 AGAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCTCAGT
48 GTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAACT
49 GTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAG
50 TTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCA
51 >962169H1       BRAITUT03        INCYTE
52 AGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTGCA
53 TTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTTTTTGAAAATAATTAAATTATCATATCTTCCATTC
54 CTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTT
55 GGGGAAATCTGAGCCTAGC
56 >4201385H1       BRAITUT29        INCYTE
57 TTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTAT
58 TAAAATTTCTGCCATTTAGAAGAAGAGAACTACATTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTATCT
59 TCACTTTATCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTATATTCTCCTTTGACATATAACTGTTGGCTTTT
60 CTAATCTGTTAAATATATCTATTTTTACCAAAGGTATTTAATAT
61 >1302516H1       PLACNOT02        INCYTE
62 AGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACCTTCTAACAGAGGAGGTAAGATTATAC
63 AGCTGCACACCTCGTAACTTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCT
64 CCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTT
65 ACTAAAAAATACCACGAGGTCC
66 >3684109H1       HEAANOT01        INCYTE
67 ATTTCATCTTCAGGATTTACAGTGCATTCTGAAANAGGAGAAATCAAACANAATTAGGAGTTGTGCAACAGCTCTTTTGA
68 GAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAANAAAATTAAATGTTGTATTAAATAGATCACCAGCTA
69 GTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGGTAATGTCAG
70 TACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTT
71 >2549720H1       LUNGTUT06        INCYTE
72 TTAGCTTGGNAAATTTTTCTAAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAA
73 AATACATGTATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGT
74 AAGTTGCAAAGACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACT
75 TATGANAGTAG
```

Fig 6

Title: long - est. mfas

```
>877279H1       LUNGAST01       INCYTE
CTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCTAAACACAATTGTTATAGCCAGAGGAACAAA
GATG ATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTGCAT
TTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGGCTTTTTGAAAATAATTAAATTATCATATCTTCCATTCC
TGTTATTGGNGG
>4713188H1       BRAIHCT01       INCYTE
CAAAGTTACTAAAAAATACCACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCACTCACCG
ACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAG
CTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGT
TTGCT
>2171082H1       ENDCNOT03       INCYTE
AGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTA
TATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCTATTTTTACCAAAGGTATTTAATATT
CTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCTAAACACAATTGTTATAGCCAGAGGAACAAA
GATGA
>875860H1       LUNGAST01       INCYTE
CTGGATTTTTCATATTTCTTATTAAAATTTCTGCCATTTAGAAGAAGAGAACTACATTCATGGTTTGGAAGAGATAAACC
TGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTATATTCTCCT
TTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCTATTTTTACCAAAGGTATTTAATATTCTTTTTTAT
GAC
>706168H1       SYNORAT04       INCYTE
GCTCATATTCACATATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGANCTATGTTGCTATGAAT
TAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAATTTCTGCCATTTAGAAGAAGAGAAC
TACATTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCANTTTATCGATAAGTCAGTTTATTTGT
TTCA
>458223H1       KERANOT01       INCYTE
ANGAGTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTT
GTTTGNTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTG
CAACAGCTCTTTTGAGAGGAGGCCTAAAGGNCAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATTAA
ATAGATC
>538436H1       LNCDNOT02       INCYTE
AAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTG
CATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTTTTTGAAAATAATTAAATTATCATATCTTCCAT
TCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTAT
>1303909H1       PLACNOT02       INCYTE
AGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACCTTCTAACAGAGGAGGTAAGATTATAC
AGCTGCACACCTCGTAACTTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCT
CCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAG
>2739211H1       OVARNOT09       INCYTE
GTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGA
GAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATTAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACG
TATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAA
GTGAGCACCTGAT
>2550343H1       LUNGTUT06       INCYTE
TGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCNAATCTTGTTAAATATATCTATTTTTACCAAAG
GTATTTAATATTCTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCTAAACACAATTGTTATAGC
CAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCATTCTCGTATGGTGCTA
>5321148H1       FIBPFEN06       INCYTE
CACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGNCAAAAATACATGTATTTCATTCTCGTA
TGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTTTTTGAAAA
TAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAAATTCAGATCCAC
CATTACTAAC
>879495H1       THYRNOT02       INCYTE
ATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAA
AGACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGT
AGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCT
TGAAAAA
>3325591H1       PTHYNOT03       INCYTE
TGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATT
AAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACG
GCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACCCTAAAGCNCC
ATGTCNNGGGCNAAAANCGAAAAAT
>543890H1       OVARNOT02       INCYTE
TTTCTAAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCA
TTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGNATAGAATTGGTAAGTTGCAAAGNCTT
TTTGAAAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGGATGGAAAATAAAAAGCAACTTATGGAAAGTAGG
ACATTCAGATC
>3733565H1       SMCCNOS01       INCYTE
CCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGNAAGANGAGACATCAAACAG
AATTAGGNGTTGTGCAAAAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTNCAATCGTGGAAAGNAAATT
AAATGTTGTATNAAATNGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGNCNGTATTCAGTCT
TTCGGAACGGCTTAGGGTAATGTCAGTACAGGANAAAAACTGTGCAGTGAG
>4641939H1       PROSTMT03       INCYTE
GTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCCATGCTGATAGGACAGACTGGAT
TTTNCATATTTCTTATTAAAATTTCTGCCATTTAGAAGAAGAGAACTACATTCATGGTTTGGNAGAGATAAACCTGAAAA
```

```
GAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTGTTTCATGTGTACATTTTTAAATATCTCCTTTGGCAT
ATAACGTGGCTTT
>200  0H1       TESTNOT03       INCYTE
TTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCTATTTTTACCAAAGGTATTTAAT
ATTCTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCTAAACACAATTGTTATAGCCAGAGGAAC
AAAGATGATATAAAATATTGTTGCTCTGANAAAAATACATGTAT
>3085331H1       HEAONOT03       INCYTE
GCTCATATTCACATATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATTTGCTATGAATT
AAACTTGTGTCGTGCTGATAGGACAGACTGGNTTTTTCATATTTCTTATTANAATTTCTGCCATTAGAAGAAGAGAACTA
CATTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTATTTCACTTTATCGATAAGTCAGT
>3414043H1       PTHYNOT04       INCYTE
GCTCATATTCACATATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAAT
TAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAATTTCTGCCATTTAGAAGAAGAGAAC
TACATTCATGGTTTGGAAGAGATAAACCTGAAA
>3705963H1       PENCNOT07       INCYTE
ANACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCCATGTCCTGGGCCTAAAATCGTATAAAA
TCTGGAnnnnnnnnnnnnnnnnnnnnGCTCATATTCACATATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTA
AAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAAT
TTCTGCCATTAGAAGAAGAGAACTACNTTCANGGTTTGGAAGAGATAACCCTGAAAAGANGGG
>5137051H1       OVARDIT04       INCYTE
AAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTNTTTGAAAATAATTAAATTATCATATCTTCCATTCCTGT
TATTGGAGATGAANATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGG
AAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCGTGCTGTNTGTCA
GTAGGAACACATCCTATTTATTGTGATGNTGTGGTTTATTAT
>3554223H1       SYNONOT01       INCYTE
ATTAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGAT
ACGGCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGGCTTAACTCTAAAG
CTCCATGTCCTGGGCCTAAAATCGTATAAAATCTGGATTTTTTTNTTTTTTTTTGCGCATATTCACATATGTAAACCAGN
ACATTCTATGTACNACAAACCTGGTTTTTAAAAAGGAAC
>4507477H1       OVARTDT01       INCYTE
GGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGGTAAT
GTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCCATGTCCTGGGCC
TAAAATCGTATAAAATCTGGA
>1955646H1       CONNNOT01       INCYTE
TGGTAAGTTGCAAAGACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGC
AACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAA
ACATAAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCGTGCTGTGCTGTGCAGTAGGGAACACATCCTATTTA
TTGTGATGTTGTGGTTTATATCCTAAACC
>4163378H1       BRSTNOT32       INCYTE
AATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGNTCTGTATTTCAGTTCCTTTCGATACG
GCTTAGGGTAATGTCAGTACAGGAAAAAAGCTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCC
ATGTCCTGGGCCTAAAATCGTATA
>5095141H1       EPIMNON05       INCYTE
AGATAAACCTGAAAAGAAGAGTGGCCTTATNTTCACTTTATCGATAAGTCAGNTTATTTGTTTCATTGTGTACATTTNNA
TATTCTCCTTTTGACATTATAACTGNTGGCTTTTCTAANCNTGTTAAATATATCTATTTTTACCAAAGGTATTTAATATT
CTTT
>9438261H1       ADRENOT03       INCYTE
TATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTTTTTGAA
AATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATG
>3451273H1       UTRSNON03       INCYTE
TTTTTTNTTTTGCTCATATTCACATATGTAAACCNGAACATTCTATGTACNACAAACCTGGTTTTTAAAAAGGAACTATG
TTGCTATGAATTAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCANATTTCTTANTAANNTTTCTGCCATTTAG
AAGA
>1402278H1       LATRTUT02       INCYTE
GTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCCATGTCCTGGGCCTAAA
ATCGTATAAAATCTGGAnnnnnnnnnnnnnnnnnnnnGCTCATATTCACATATGTAAACCAGAACATTCTATGTACTACAAA
CCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCATATTT
CTTA
>4361191H1       SKINNOT01       INCYTE
GCAAAGACTTTTTGANAATNATTAANTTATCATATCTTCCATTCCTGTTATNGGAGATGANAATAAAAAGCAACTTATGA
AAGTAGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCNCAGAAAAACATAAAGC
ACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCGTGCTGTGCTGTGCAGTAGGAACACATCCNATTTATTGTGNTGTN
GNGGTTTTATGATC
>1307017H1       PLACNOT02       INCYTE
TGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCCATGTCCTGGGC
CTAAAATCGTATAAAATCTGGAnnnnnnnnnnnnnnnnnnnnGCTCATATTCACATATGTAAACCAGAACATTCTATGTACT
ACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCATGCTGATAGGACAGACTGGATTTTTCA
TAT
>5032225H1       HEARFET03       INCYTE
AATTATCATATCTTCCATTCCTGTTATTGGAGATGNAAATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCAT
TACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCTTGAAAAAGACTGTCAGCTTC
CTGATAAAGCGTGCTGTGCTGTGCAGTAGGAACACATCCTATTTATTGTGATGTTGTGGTTTTATTATCTTAAACTCGTT
CCAT
>3732621H1       SMCCNOS01       INCYTE
ANAGATGATATAAAANATTGTTGCTCTGACAANNATACATGTATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTG
```

>3530274H1 BLADNOT09 INCYTE
TTCC[illegible]TCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTATT
CCTT[illegible]TGGGGAAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCG
TGCTGTGCTGTGCAGTAGGAACACATCCTATTTATTGTGATGTTGTGGTTTTATTATCTAAACTCTGTTCCATACACTTG
TATAAATACATGGATATTTTTATGTACAGAAGTATGTCTCTTAACCAGTTCA
>3530249H1 BLADNOT09 INCYTE
CTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGANAGTAGACATTCAGATCCAGCCATTACTAACCTAT
TCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGC
GTGCTGTGCTGTGCAGTAGGAACACATCCTATTTATTGTGATGTTGTGGTTTTATTATCTTAAACTCTGTTCCATACACT
TGTATAAATACATGGATATTTTTATGTACAGAAGTATGTCTCTTAACCAGTTCACTTATTGTACCTGG

Fig 6 (cont'd)

VEGFE1	AAAATGTATGGATACAACTTAC	22
VEGFE2	GTTTGATGAAAGATTTGGGCTTG	23
VEGFE3	TTTCTAAAGGAAATCAAATTAG	22
VEGFE4	GATAAGATTTGTATCTGATG	20
VEGFE5	GATGTCTCCTCTTTCAG	17
VEGFE6	GCACAACTCCTAATTCTG	18
VEGFE7	AGCACCTGATTCCGTTGC	19
VEGFE8	TAGTACATAGAATGTTCTGG	20
VEGFE9	AAGAGACATACTTCTGTAC	19
VEGFE10	CCAGGTACAATAAGTGAACTG	21

Fig. 7

Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure

Receipt in the case of an original deposit issued pursuant to Rule 7.1 by the International Depositary Authority BCCM™/LMBP identified at the bottom of next page

***International* Form BCCM™/LMBP/BP/4/99-23**

To : **Name of the depositor :** Janssen Pharmaceutica N.V.

Address : Turnhoutseweg 30
B-2340 Beerse
Belgium

I. Identification of the microorganism:

I.1 Identification reference given by the depositor:

VEGF-X CUB PET22b

I.2 Accession number given by the International Depositary Authority:

LMBP 3991

II. Scientific description and/ r proposed tax n mic designation

The microorganism identified under I above was accompanied by:

(mark with a cross the applicable box(es))

– a scientific description	yes ☒	no ☐
– a proposed taxonomic designation	yes ☐	no ☒

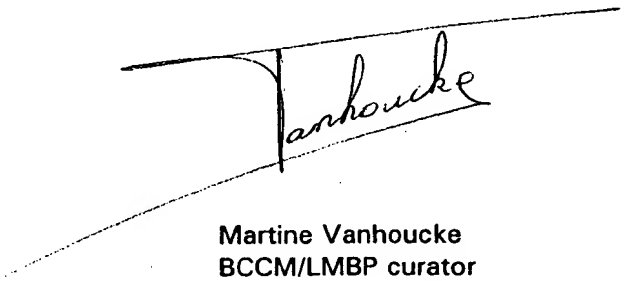
III. Receipt and acceptance

This International Depositary Authority accepts the microorganism identified under I above, which was received by it on (date of original deposit) : December 20, 1999

IV. International Depositary Authority

Belgian Coordinated Collections of Microorganisms (BCCM™)
Laboratorium voor Moleculaire Biologie - Plasmidencollectie (LMBP)
Universiteit Gent
K.L. Ledeganckstraat 35
B-9000 Gent, Belgium

Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s):



Date : January 12, 2000

Martine Vanhoucke
BCCM/LMBP curator